EM1438

Re submit 3/12/04

Document And Report Documentation Page Submitted as edoc_1075483867

Report Documentation Page	*Form Approved OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
22 JUL 2003	**N/A**	**-**

4. TITLE AND SUBTITLE	
Algebraic Equivalence of Conjugate Direction and Multistage Wiener Filters	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Office of Naval Research Supported by, CSU, MRC, SAIC**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
Also see: ADM001520, ASAP Workshop 2003 , The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:	17.	18.	19a. NAME OF RESPONSIBLE

a. REPORT	b. ABSTRACT	c. THIS PAGE	LIMITATION OF ABSTRACT	NUMBER OF PAGES	PERSON
unclassified	**unclassified**	**unclassified**	**UU**	**13**	**Patricia Mawby, EM 1438** **PHONE:(703) 767-9038** **EMAIL:pmawby@dtic.mil**

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

pwd: cannot determine current directory!

Colorado State University

Algebraic Equivalence of Conjugate Direction and Multistage Wiener Filters

Louis Scharf , Todd McWhorter , Edwin Chong ,
Scott Goldstein , Michael Zoltowski

Supported by Office of Naval Research under various contracts to
CSU , MRC , SAIC

Program:

Define a very general class of *iterative subspace* Wiener filters. Within this class, identify two interesting subclasses: *conjugate direction* Wiener filters and *multistage* Wiener filters. Study their equivalences.

Key Findings

1. **Any subspace expansion in an iterative subspace Wiener filter that uses the direct sum of the past subspace and the current gradient turns out an orthogonal basis for the Krylov subspace.**

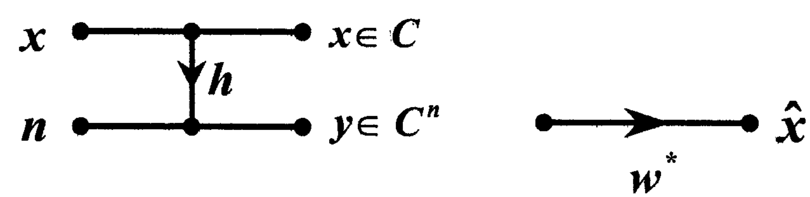
2. **For every multistage Wiener filter, there is a family of equivalent conjugate direction Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.**

3. **For every *orthogonal* multistage Wiener filter, there is a family of equivalent conjugate *gradient* Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.**

4. **If an *orthogonal* multistage Wiener filter is initialized at the cross-covariance, then it and its equivalent conjugate *gradient* Wiener filter turn out a Krylov subspace. And vice-versa.**

A Few (Among Many) Motivating Problems

1. General linear channel



2. Repeated measurements

$$y_i = x + n_i \quad \& \quad y = hx + n\,;\ h^T = [1 \quad 1 \quad 6 \quad 1]\ \text{: averaging vector}$$

3. Multisensor Array Processing

$$y_i = e^{j\rho i}x + n_i \ \& \ y = hx + n\,;\ h^T = [1 \quad e^{j\rho} \quad 6 \quad e^{j\rho(n-1)}]\ \text{: steering vector}$$

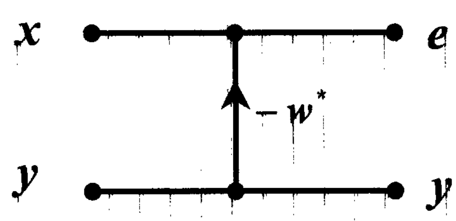
4. CDMA

$$y_i = h_i x + n_i \quad \& \quad y = hx + n\,;\ h^T = [h_1 \quad 6 \quad h_n]\ \text{: chipping vector}$$

Generally, x is an unobserved complex scalar and y is a measured vector that carries information about x, linearly.

Second-Order Characterization of The Two Channel Filtering Problem

$$\begin{array}{c} x \cdot \\ y \cdot \end{array} \begin{bmatrix} r_{xx} & r^* \\ r & R \end{bmatrix} \; ; r_{xx} = Exx^* \in \Re^+ \;\; \& \;\; r = Eyx^* \in C^n \;\; \& \;\; R = Eyy^* \in C^{n\times n}$$



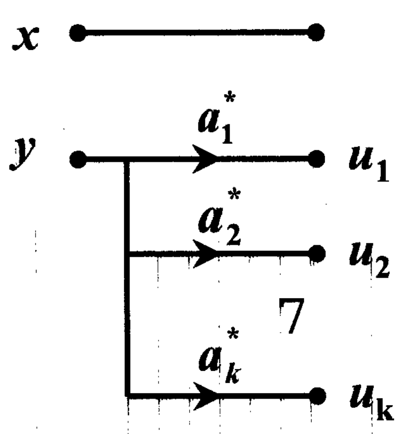
$$\begin{bmatrix} Q_{xx}(w) & -\gamma^*(w) \\ -\gamma(w) & R \end{bmatrix} \; ; \quad Q_{xx}(w) = r_{xx} - r^* w - w^* r + w^* R w \;\; : \text{MSE}$$

$$\gamma(w) = Rw - r : \quad \textit{gradient}$$

For $\quad w = R^{-1} r, \quad \begin{bmatrix} r_{xx} - r^* R^{-1} r & 0 \\ 0 & R \end{bmatrix}.$

Thus, the plan is always to take the cov(x, y) to block diagonal form, where zeros reveal the orthogonality principle at work ($\gamma(w) = 0$).

The Expanding Subspace Idea: Resolve onto a Basis



$$\begin{matrix} x \cdot \\ u \cdot \end{matrix} \begin{bmatrix} r_{xx} & r^* A_k \\ A_k^* r & A_k^* R A_k \end{bmatrix}; \quad A_k = [a_1 \quad a_2 \quad 6 \quad a_k] \in C^{n \times k}$$

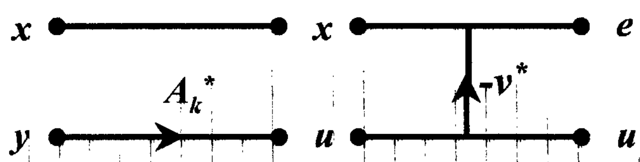
: direction vectors or filters or coordinate vectors or basis vectors.

Think of this as a subspace expansion until the right tradeoff is achieved between bias and variance … in advance of detection, estimation, beamforming, spectrum analysis.

Generalize to matrix a_i for vector x. We hope it generalizes to (true) filters $a_i(z)$ for time series.

Filtering in the Expanding Subspace

The original (x, y) problem is now an (x, u) problem:



$$\begin{bmatrix} Q_{xx}(A_k v) & -\gamma^*(A_k v)A_k \\ A_k^*\gamma(A_k v) & A_k^* R A_k \end{bmatrix}$$

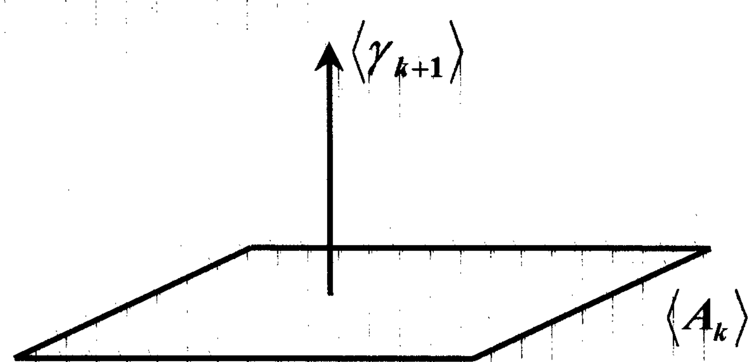
We are stuck with A_k, but we can optimize v. For the Wiener solution $v_k = (A_k^* R A_k)^{-1} A_k^* r \quad (w_k = A_k v_k)$, the covariance is

$$\begin{bmatrix} r_{xx} - r^* A_k (A_k^* R A_k)^{-1} A_k^* r & 0 \\ 0 & A_k^* R A_k \end{bmatrix} ; \quad 0 = A_k^* \gamma(A_k v_k)$$

This bears comment.

Properties of the Expanding Subspace $\langle A_k \rangle$

In this construction call $A_k = [a_1 \quad a_2 \quad 5 \quad a_k]$ a basis for the k-dimensional subspace $\langle A_k \rangle$ and $u = A_k^* y$ the coordinates of y in this basis. Call $\gamma_{k+1} = \gamma(A_k v_k)$ the gradient of $Q_{xx}(w)$ at the *subspace* solution $w_k = A_k v_k$, and $\Gamma_k = [\gamma_1 \quad \gamma_2 \quad 5 \quad \gamma_k]$ the matrix of gradients. Then $A_k^* \gamma_{k+1} = 0$ or $A_k^* \Gamma_k = C_k$: lower triangular … all a consequence of principle of orthogonality. The picture is this:



If $\langle A_k \rangle$ is initialized at $A_1 = a_1 = -\gamma_1 = r - Rw_0 = r$, then $[A_k, \gamma_{k+1}]$ turns out an orthogonal basis for the Krylov subspace $\langle K_{k+1} \rangle = \langle r, Rr\ ,\ 5\ ,\ R^k r \rangle$.

Conjugate Directions and Multistage Wiener Filters

So far we have required nothing special of the filters or direction vectors A_k. So let us now impose constraints

CDWF	MSWF
$A_k = D_k$: *re-name*	$A_k = G_k$: *re-name*
$D_k^* R D_k = \Sigma_k^2$: *diagonal*	$G_k^* R G_k = T_k = B_k \Sigma_k^2 B_k^*$: *tri-diagonal*
$D_k^* \Gamma_k = C_k$: *lower triangular*	$G_k^* \Gamma_k = B_k C_k$: *lower triangular*
$(\langle D_k \rangle \text{ not generally Krylov})$	$(\langle G_k \rangle \text{ not generally Krylov})$

$$\text{CD} \longleftrightarrow \text{MS}$$

$$D_k B_k^* = G_k \quad \text{or} \quad d_k b_{kk} + d_{k-1} b_{k-1k} = g_k \quad : \quad AR$$

By constructing AR recursion between direction vectors, we go back and forth between CD and MS implementations.

Conjugate Gradient & Orthogonal Multistage Wiener Filter

CGWF	OMSWF
$A_k = D_k$: $re-name$	$A_k = G_k$: $re-name$
$D_k^* R D_k = \Sigma_k^2$: $diagonal$	$G_k^* R G_k = T_k = B_k \Sigma_k^2 B_k^*$: $tri-diagonal$
$D_k^* \Gamma_k = C_k$: $lower\ triangular$	$G_k^* \Gamma_k = B_k C_k$: $lower\ triangular$
$D_k B_k^* = \Gamma_k$: $conjugate\ gradient$	$G_k^* G_k = diagonal\ \ (orthogonal)$
$(\langle D_k \rangle\ generally\ \ Krylov)$	$(\langle G_k \rangle\ generally\ \ Krylov)$

$\Gamma_k \Lambda_k = G_k$: CG $\longleftrightarrow$ OMS : $G_k = \Gamma_k$

$D_k B_k^* = \Gamma_k$ or G_k : AR

By constructing AR recursion between direction vectors, we go back and forth between CG and OMS implementations.

A Little More Detail about Implementation of CD and MS Wiener Filters

In the analysis stage, the order in which zeros are forced into the covariance matrices is this

$$\begin{bmatrix} r_{xx} & r^T \\ r & R \end{bmatrix} \xrightarrow[A^* = D^*]{\text{CD}} \begin{bmatrix} r_{xx} & s_1 & s_2 & s_3 & s_4 \\ s_1 & \sigma_1^2 & \boxed{1} & \boxed{2} & \boxed{3} \\ s_2 & \boxed{1} & \sigma_2^2 & \boxed{2} & \boxed{3} \\ s_3 & \boxed{2} & \boxed{2} & \sigma_3^2 & \boxed{3} \\ s_4 & \boxed{3} & \boxed{3} & \boxed{3} & \sigma_4^2 \end{bmatrix} \quad : \textit{diag.}$$

$$\begin{bmatrix} r_{xx} & r^T \\ r & R \end{bmatrix} \xrightarrow[A^* = G^*]{\text{MSWF}} \begin{bmatrix} r_{xx} & \delta_1 & \boxed{1} & \boxed{1} & \boxed{1} \\ \delta_1 & \gamma_1^2 & \delta_2 & \boxed{2} & \boxed{2} \\ \boxed{1} & \delta_2 & \gamma_2^2 & \delta_3 & \boxed{3} \\ \boxed{1} & \boxed{2} & \delta_3 & \gamma_3^2 & \delta_4 \\ \boxed{1} & \boxed{2} & \boxed{3} & \delta_4 & \gamma_4^2 \end{bmatrix} \quad : \textit{tri-diag.}$$

The zero pattern of CD makes direction vectors *R-conjugate*. The zero pattern of MSWF makes direction vectors *R-tridiagonal*. In the case of CGWF and OMSWF, the *multistage vectors* of the OMSWF are identical to the *gradient vectors* of the CGWF.

In their synthesis steps CDWF and MSWF take their analyzed covariance matrices to diagonal form, where the orthogonality condition is enforced. Therefore, they have the same minimum MSE:

$$\left[\begin{array}{c|cccc} r_{xx} & & \mathbf{s}^T & & \\ \hline & \sigma_1^2 & 0 & 6 & 0 \\ \mathbf{s} & 0 & \sigma_2^2 & 0 & 7 \\ & 7 & 0 & \sigma_3^2 & 0 \\ & 0 & 6 & 0 & \sigma_4^2 \end{array}\right] \xrightarrow{\textbf{CD}} \left[\begin{array}{c|cccc} Q_{xx} & & \mathbf{0}^T : \text{ortho} & & \\ \hline & \sigma_1^2 & 0 & 6 & 0 \\ \mathbf{0} & 0 & \sigma_2^2 & 0 & 7 \\ & 7 & 0 & \sigma_3^2 & 0 \\ & 0 & 6 & 0 & \sigma_4^2 \end{array}\right]$$

$$\begin{bmatrix} r_{xx} & \delta_1 & 0 & 6 & 0 \\ \delta_1 & \gamma_1^2 & \delta_2 & 0 & 7 \\ 0 & \delta_2 & \gamma_2^2 & \delta_3 & 0 \\ 7 & 0 & \delta_3 & \gamma_3^2 & \delta_4 \\ 0 & 6 & 0 & \delta_4 & \gamma_4^2 \end{bmatrix} \xrightarrow{\textbf{MSWF}} \left[\begin{array}{c|cccc} Q_{xx} & & \mathbf{0}^T : \text{ortho} & & \\ \hline & \hat{\gamma}_1^2 & 0 & 6 & 0 \\ \mathbf{0} & 0 & \hat{\gamma}_2^2 & 0 & 7 \\ & 7 & 0 & \hat{\gamma}_3^2 & 0 \\ & 0 & 6 & 0 & \hat{\gamma}_4^2 \end{array}\right]$$

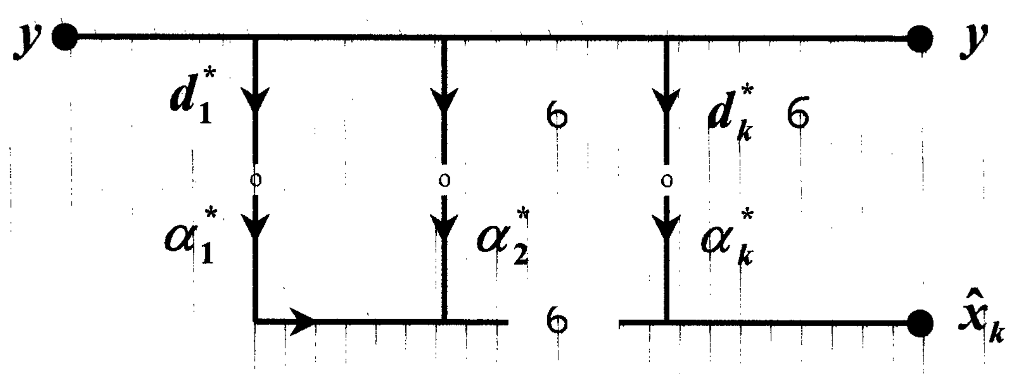
This proves equivalence of CDWF and MSWF. Their respective MMSE formulas are

$$Q_{xx} = r_{xx} - \frac{s_1^2}{\sigma_1^2} - 6 - \frac{s_k^2}{\sigma_k^2} = r_{xx} - \cfrac{\delta_1^2}{\gamma_1^2 - \cfrac{\delta_2^2}{\gamma_2^2 - \cfrac{\delta_3^2}{5 - \gamma_k^2}}}$$

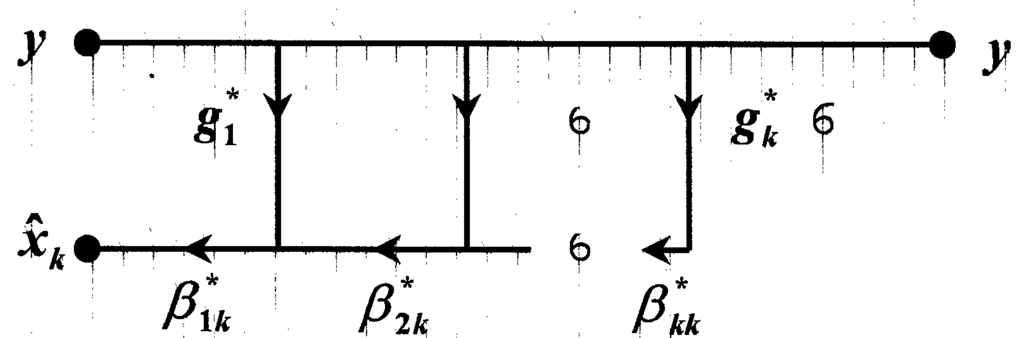
: continued sum vs. continued fraction

The picture is this.

CDWF:



MSWF:



For an optimization theorist, a *line search* for the optimum *step size* α_i or β_i in direction d_i or g_i is equivalent to the filtering theorist's MMSE *weight computation* in the synthesis stage of the iterative filter.

Conclusions

1. Any subspace expansion in an iterative subspace Wiener filter that uses the direct sum of the past subspace and the current gradient turns out an orthogonal basis for the Krylov subspace.
2. For every multistage Wiener filter, there is a family of equivalent conjugate direction Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.
3. For every *orthogonal* multistage Wiener filter, there is a family of equivalent conjugate *gradient* Wiener filters, with identical subspaces, gradients, and mean-squared errors. And vice-versa.
4. If an *orthogonal* multistage Wiener filter is initialized at the cross-covariance, then it and its equivalent conjugate *gradient* Wiener filter turn out a Krylov subspace. And vice-versa.
5. All of this generalizes to matrix filters a_i. Does it generalize to time series filters $a_i(z)$? To something like least squares filtering on manifolds, as opposed to (linear) subspaces? To kernel-based nonlinear processing of measurements?